The effect of social preferences on the evolution of cooperation in public good games

Marco A. Janssen
Center for the Study of Institutional Diversity
School of Human Evolution and Social Change
Arizona State University
PO Box 872402
Tempe, AZ 85287-2402
Phone: 480 727 7067
Fax: 480 965 7671
Marco.Janssen@asu.edu

Miles Manning and Oyita Udiani
Mathematical, Computational and Modeling Sciences Center
Center for the Study of Institutional Diversity
School of Human Evolution and Social Change
Arizona State University

Abstract
Human societies are unique in the level of cooperation among non-kin. Evolutionary models explaining this behavior typically assume pure strategies of cooperation and defection. Behavioral experiments, however, demonstrate that humans are typically conditional co-operators who have other-regarding preferences. Building on existing models on the evolution of cooperation and costly punishment, we use a utilitarian formulation of agent decision making to explore conditions that support the emergence of cooperative behavior. Our results indicate that cooperation levels are significantly lower for larger groups in contrast to the original pure strategy model. Here, defection behavior not only diminishes the public good, but also affects the expectations of group members leading conditional co-operators to change their strategies. Hence defection has a more damaging effect when decisions are based on expectations and not only pure strategies.

Key Words: Public good games, group selection, other-regarding preferences, conditional cooperation

Introduction

One of the key puzzles in social and biological sciences is the scale of cooperative behavior in human societies. Unlike other species, humans cooperate in non-repeated interactions with non-kin in large groups. Theories of cooperation have focused on kin selection [18], direct reciprocity in repeated interactions ([3]; [31]), and indirect reciprocity and costly signalling of reputation ([5]; [24]; [33]). The structure of reciprocity interactions in spatial networks and group selection have also been proposed as possible explanations [23].

Most computational and mathematical models on the emergence of cooperation have employed evolutionary game theory with pure or fixed mixed strategists to capture human behavior ([4]; [16]; [23]). However, experimental studies show that humans make strategic decisions that are contextual and evolve due to a learning process [9]. Participants in behavioral experiments have been found to adjust their level of cooperation in response to an expected level of cooperation by others (e.g., [15]). Thus, participant decisions made in social dilemma situations cannot be explained by selfish rational motivations. Alternative utility functions that have been proposed include weighting the earnings of others relative to individual returns ([10], [13]). As such, participants may not maximize their material wealth when maximizing their utility. Finally, decisions are typically made with partial (or erroneous) information about the intended actions of others. In iterated games, such uncertainty may be mitigated by using prior and current information to inform the expected decisions in future rounds [9].

The use of punishment has been found to be an important factor explaining cooperation in large human societies. Laboratory experiments on public goods and common pool resources have shown that participants are willing to give up monetary returns to punish non-cooperators ([11]; [14]; [28]; [32]). When the option to punish others at a cost to themselves was introduced in experiments, participants utilized this force and the level of cooperation increased. From an evolutionary perspective, it is puzzling why individuals would accept a reduction in payoff to decrease the payoff of other individuals. The possibility of the evolution of costly punishment in human societies has been demonstrated by various studies including Boyd et al. [7].

Laboratory experiments with public good games show that participants invest initially around 50% of their endowment in linear public good games and typically reduce their investments in subsequent rounds [11]. When costly punishment is allowed, participants use punishment and the level of contributions to the public good goes up. The question addressed in this paper is not to model decision making of those laboratory experiments itself, but to use an evolutionary model to explore the conditions that enable humans to evolve who have other regarding preferences as proposed as potential explanations for the observed behavior in experiments. More precisely, can we evolve agents who make decisions as we observe in behavioural experiments?

In order to do this we model agent decision making in line with experimental evidence of human behavior in social dilemmas. Typical models that study the evolution of cooperation assume agents being co-operators or defectors [24]. In contrast those previous models on the evolution of cooperation, we assume that the trade-offs between alternative behavioral actions also depend on an expected utility, which includes “other-regarding preferences” (i.e., considerations for the earnings of other group members). Finally, we incorporate a strategy that allows agents to weight future expectations based on the current level of cooperation in the group.

With these modifications, we explore some conditions that favour the evolution of preferences and expectations that lead to cooperative behavior. To achieve this, we will build on the cultural group selection model of Boyd et al. [7] which was used to evaluate the conditions for the evolution of costly punishment. As such we will also test the conditions under which cooperative behavior evolve with and without the option of punishment.

Recent work on the evolution of preferences by Alger and Weibull [1] suggests that the emergence of individuals who value the group payoffs is possible conditional on a sufficient degree of assortativity in the pair-wise interactions within an infinite population. This paper uses a specific form of assortive interactions (cultural group selection) in games with n (>2) individuals, and we show that group size affect the level of cooperation. Furthermore, we look at the conditions of costly punishment evolving.

The rest of the paper will be structured as follows: First, we present the model in detail and its underlying assumptions. Next, we present the results of our analysis and simulations, and conclude with a discussion including implications for future work. The model code and detailed documentation of the model used in this paper can be found at http://www.openabm.org/model/3887/version/2/view.

Model Description

Consider a population that is divided into groups of size n. Agents make a decision to contribute or not to the public good. Agents have an initial endowment of 1 unit, and if an agent contributes it will incur a cost c to produce a total benefit b that is shared equally among group members. If an agent does not contribute, it will incur no costs and produce no benefits. If the fraction of cooperators in the group is y, the expected payoff x for cooperators is $1 + by - c$ and the expected payoff for non-cooperators is $1 + by$. Hence the payoff disadvantage of the cooperators is a constant *c* independent of the distribution of types in the population. After all agents make their decisions a new generation of agents is created based on relative payoffs.

An agent decides whether or not to contribute to the public good of her group based on the evaluation of the expected utility of two options. In line with the indirect evolutionary approach individuals make decisions based on the subjective utility, but the selection of offspring is based on the objective payoff [17]. Social preferences are proposed to explain observations in social dilemma experiments where participants may weigh the earnings of others in additional to individual welfare. There are two types of social preference models: (1) Distributional preferences assume people care about the differences of payoffs among participants and are motivated to reduce differences (e.g. [13]) or help those who have the lowest payoffs (e.g. [2]). (2) Reciprocity preferences models which assume people have perceived norms of fairness and respond to participants contributions based on these beliefs (e.g. [29]). Empirical analysis that aims to test which models of social preferences best explain the data conclude that we cannot find one particular model that fits all participants [10].

Charness and Rabin [10] proposes a comprehensive model of individual utility that captures different models of social preferences for two-player games with perfect information. Although they apply the utility function to two-player games, we apply this to games of larger number of players assuming that the agents only have information of their own earnings and the average earnings of the others. Charness and Rabin [10] also discuss a version of the model with reciprocal preferences, but we will – for simplicity's sake – only focus on the different types of distributional preferences. Therefore, agents decide to contribute or not in a public good game using the following utility function (U_i):

$$U_i = \alpha_i * \bar{x}_{-i} + (1 - \alpha_i) * x_i \; if \; x_i > \bar{x}_{-i} \quad (1a)$$

$$U_i = \beta_i * \bar{x}_{-i} + (1 - \beta_i) * x_i \; if \; x_i < \bar{x}_{-i} \quad (1b)$$

Here, x_i defines the earning of the focal agent and $\bar{x}_{-i}$ the average earning of all other agents in the group. The parameter α_i defines the weight for the average earnings of others compared to the earnings of oneself if one has higher earnings than others. Similarly, the parameter β_i defines the weight for the earnings of others if one is earning less than the others. Charness and Rabin [10] assume $\beta \leq \alpha$ which means that the preference for gains relative to others is at least as high when having less as when having more. Together, these parameters, along with expected contribution of the group into the public good, determine whether an agent will cooperate. Note that the decision to contribute may depend not only on individual payoffs, but also on the payoffs of others. In line with Charness and Rabin [10], we can define the following cases of distributional preferences for$\beta_i \leq \alpha_i \leq 1$.

Case 1: The players prefer to have their payoffs higher than those of the other players.

If $\beta_i \leq \alpha_i \leq 0$, players are highly competitive.

Case 2: Players prefer the payoffs among all players to be equal. This "inequity aversion" holds when $\beta_i < 0 < \alpha_i \leq 1$ (see [12]).

Case 3: The third model captures the Andreoni and Miller [2] model where participants prefer to get more for all players, but prefer to get more for themselves when they have lower earnings than the average. This is presented when $0 < \beta_i \leq \alpha_i \leq 1$.

Case 4: If $\alpha_i = \beta_i = 0$, then players only care about their own welfare.

Agents make these calculations based on an expectation of cooperation within their group. This "trust" variable (T_c) has an initial value and is updated based on observed levels of cooperation by other group members after each round of play.

Now we must define the expected earnings on which the agent decides to contribute or not to the public good. In order for agents to make decisions, they need to evaluate the expected utility value of equation (1). Therefore, we need to define the expected earnings, which is based on the expected level of cooperation. For a cooperative agent the expected fraction of cooperation for by (n-1) other group members is T_c and 1 for the agent itself. Therefore we can define the expected earnings for a cooperative agent as

$$x_C^* = 1 + b \cdot \frac{\{T_C \cdot (n-1) + 1\}}{n} - c \tag{2}$$

In a similar way we can define the expected earning of a defecting agent as

$$x_D^* = 1 + b \cdot \frac{\{T_C \cdot (n-1)\}}{n} \tag{3}$$

In order to calculate the utility of agent i, agent i needs to define the expected average earning of the other n-1 agents. We can distinguish the expectations for the case agent i will contribute or not contribute. If agent i contributes, conditional to the expectation that fraction T_c of the other agents contributes too, the expected earnings of other agents is equal to

$$\overline{x_C^*} = T_c \cdot x_C^* + (1 - T_c) \cdot \left(x_D^* + \frac{b}{n}\right) \tag{4}$$

Conversely, if agent i does not contribute, the benefit b/n is not enjoyed by other agents. Thus, the expected earnings of other agents is

$$\overline{x_D^*} = T_c \cdot \left(x_C^* - \frac{b}{n}\right) + (1 - T_c) \cdot x_D^* \tag{5}$$

Using equation (1) we can define the expected utility of cooperation and defection in the following way:

$$E[U_C] = \alpha_i * \overline{x_C^*} + (1 - \alpha_i) * x_C^* \;\; if x_C^* > \overline{x_C^*} \tag{6a}$$

$$E[U_C] = \beta_i * \overline{x_C^*} + (1 - \beta_i) * x_C^* \;\; if x_C^* < \overline{x_C^*} \tag{6b}$$

$$E[U_D] = \alpha_i * \overline{x_D^*} + (1 - \alpha_i) * x_D^* \;\; if x_D^* > \overline{x_D^*} \tag{7a}$$

$$E[U_D] = \beta_i * \overline{x_D^*} + (1 - \beta_i) * x_D^* \;\; if x_D^* < \overline{x_D^*} \tag{7b}$$

The expected utility values are used to make a decision. The choice with the highest expected utility is chosen. If there is a tie, randomly one of the options is drawn. With a small probability, ζ, an error is made and the alternative is selected. If there is a tie one of the options is drawn randomly.

When we allow agents to punish others at a cost to themselves, the cost of punishment is *k* and the penalty of the agent being punished is *p*. This will affect the expected earnings of the agents. Empirical studies have shown that humans can participate into anti-social punishment and retaliation ([12]; [20]) and this could be included into models on the evolution of punishment (e.g. [19]; [30]). For simplicity's sake, we assume only cooperative agents will use costly punishment. The expected earnings of a cooperative agent who is punishing is therefore defined as:

$$x_{CP}^{*} = 1 + b \cdot \frac{\{T_C \cdot (n-1)+1\}}{n} - c - k(1 - T_c) \tag{8}$$

The expected earning of a defecting agent (3) is now updated as follows:

$$x_{D}^{*} = 1 + b \cdot \frac{\{T \cdot (n-1)\}}{n} - p \cdot T_p \cdot T_c \tag{9}$$

With T_p the share of cooperative agents who punish defectors.

The expected average earning of other agents can now be updated for the three possible decisions of agent i. If agent i is contributing to the public good, but does not punish, the expected average earnings of the other agents is equal to

$$\overline{x_{C}^{*}} = T_c \cdot \{(1 - T_p) \cdot x_{C}^{*} + T_p \cdot x_{CP}^{*}\} + (1 - T_c) \cdot \left(x_{D}^{*} + \frac{b}{n}\right) \tag{10}$$

The expected earning of other agents when agent i is a punisher is

$$\overline{x_{CP}^{*}} = T_c \cdot \{(1 - T_p) \cdot x_C^{*} + T_p \cdot x_{CP}^{*}\} + (1 - T_c) \cdot \left(x_D^{*} + \frac{b}{n} - p\right) \tag{11}$$

The expected earning of other agents when agent i is defecting is given by

$$\overline{x_{D}^{*}} = T_c \cdot \{(1 - T_p) \cdot \left(x_C^{*} - \frac{b}{n}\right) + T_p \cdot \left(x_{CP}^{*} - \frac{b}{n} - k\right)\} + (1 - T_c) \cdot x_D^{*} \tag{12}$$

The expected utility of a cooperative agent who uses costly punishment is now defined as

$$E[U_P] = \alpha_i * \overline{x_{CP}^{*}} + (1 - \alpha_i) * x_{CP}^{*} \;\; if x_{CP}^{*} > \overline{x_{CP}^{*}} \tag{13a}$$

$$E[U_P] = \beta_i * \overline{x_{CP}^{*}} + (1 - \beta_i) * x_{CP}^{*} \;\; if x_{CP}^{*} < \overline{x_{CP}^{*}} \tag{13b}$$

If an agent has decided to cooperate, the agent also has to make a decision to punish or not. The decision to punish or not is based on the expected utility of using punishment versus no punishment.

The agent evaluates the expected utility of both cooperation and defection and chooses the option with the highest expected utility. If there are multiple options with equal utility, one of the options is chosen at random. Like Boyd et al.[7], we assume a small probability ζ the agent will make a mistake. When an agent cooperates she evaluates the expected utility to punish or not, and chooses the option with the highest expected utility.

Every generation agents make decisions, and transfer the expectations about other agents to the next generation. This means that they update the expectations that others cooperate (T_c) and the expectation that other cooperator punish (T_p) based on the observations made during the last generation.

Every generation agents may update the behavioral norms (i.e., α, β) based on observations of a more successful agent, meaning an agent with higher earnings. With probability (1-*m*) this agent is from the same group, and with probability m the agent is from another group in the population.

This models the individual-level selection of forces that promote payoff-maximizing strategies. This submodel is run with probability m for each agent. Agent i copies the behavioral traits of agent *j* with a probability equal to $\frac{\left(1+\left(X_j-X_i\right)\right)}{2}$.

Like Boyd et al. [7] we assume that group selection occurs through intergroup conflicts. At the end of each generation, groups are randomly paired, and with probability ε there is a conflict and then the interaction results in one group defeating and replacing the other group. The probability that group *i* defeats group *j* is $\frac{\left(1+\left(N_{c,j}-N_{c,i}\right)\right)}{2}$, where $N_{c,l}$ is a fraction of cooperators in group l. This assumes agents who are more successful in generating the public good for their group, are more likely to be imitated. As a consequence, cooperation is the sole target of the resulting group selection process.

Finally, every generation a mutation process is implemented. New values for α, β, T_c and T_p are drawn from normal distributions with mean equal to the current value and standard deviation μ. Table 1 summarizes all variables used in the model.

Analysis

Before presenting our results, we will discuss the decisions agents can make based on different levels of trust, and different values of α and β. We can derive for each group size the combinations of α and β that lead to the emergence of cooperation or not under fairly general conditions (See appendix). Figure 1 shows that for higher levels of trust agents are contributing with lower levels of β. Thus with higher levels of trust, agents who earn less than average still contribute. If trust decreases and agents earn less than the group average, only agents who focus on the average earnings of others will contribute. Hence we can expect that a loss of trust will lead to more agents free-riding, which further erodes trust. When group size is larger, we see that the range of agent types that are willing to contribute decreases. Thus for larger groups we will expect to see less cooperation evolving.

When punishment is allowed we see that only agents who have negative α and β are willing to use punishment if the level of expected cooperation and punishment is high enough (Figure 2). This suggests that the evolution of costly peer-punishment leads to highly competitive agents, as has been suggested by Jensen [21].

Our main interest was to explore the consequences of the inclusion of agents who make decisions based on expected utility, when that utility includes other-regarding preferences. The model was run 50 times for group sizes 2, 4, 8, 16, 32, 64, and 128 agents. Each model run consisted of 128 groups, and was simulated for 5000 generations with relevant metrics (e.g., fraction cooperated/punish) calculated at the end of the run as the average of the last 1000 generations. We look only at the last 1000 generations to focus on the evolved behavior of the agents which stabilizes after a few 1000 generations (Figures 3 and 4). We used the following default values $c = k = 0.2$; $p = 0.8$; b=0.5; m=0.01; ε=0.015; and μ=0.01 in line with the values used by Boyd et al. [7]. The initial values for α, β, T_c and T_p are zero, representing the selfish rational actors. However, one of the 128 groups consisted initially of cooperative agents, again in line with Boyd et al. [7] (i.e., $\alpha=\beta=T_c=T_p=1$).

To illustrate the dynamics of the model Figures 3 and 4 show typical simulations for group size 8 with and without punishment. Without the option of punishment we see that the level of cooperation remains low for about 2000 generations, before sufficient groups start cooperating and cooperation spreads quickly through the population due to group selection (Figure 3). Agents are initially selfish before pro-social preferences evolve after around 2000 generations (Figure 4). When punishment is allowed, the dynamics are quite different. Due to mutation and imitation cooperative behaviour can evolve and groups with cooperative agents who are willing to punish can enforce cooperation (Figure 3). T_P evolves rapidly to a substantial level. But a high T_P does not mean a higher degree of punishment since punishment only happens if a member of the group does not contribute. Initially, between group selection forces favour groups with pro-social traits and thus explain the high level of cooperation (Figure 4). However, once cooperation is established, within group forces favor competitive agents who can exploit defective behavior. Consequently, competitive phenotypes are selected for the long term.

Next, we present summary statistics for 50 runs of different group sizes. As expected we find a lower level of cooperation in larger groups (Figure 5). This is similar as the original Boyd et al. [7] model with pure strategies. One difference between the original Boyd model and conditional cooperation model is the steeper decline of cooperation with larger group sizes. For small group sizes, the level of cooperation with conditional cooperation is higher than the model with the pure strategies (Boyd model). When group size is 16 (no punish) or 32 (punish) the conditional cooperation model experienced a sharp drop in cooperation. The rapid decline of cooperation in larger groups can be explained by the interaction of trust and defection. If an agent defects in a group, other agents reduce the trust in their group members as trust is based on observed cooperation levels, and Figure 1 shows that more agent types will start to defect too. Thus without a change of the behavioral attributes of the agent, the observed behavior changes due to changes in expectations. Our model demonstrates that the frequency of cooperative outcomes might change due to both imitation of successful strategies as well as changes in expectations based on prior experiences.

When punishment is allowed there is an increased level of cooperation for larger group sizes. Still if group size is 32 or more the level of cooperation remains low (Figure 5). Figure 2 provides some intuition for this observation. When groups are large, only highly pro-social agents are willing to cooperate if trust is low. It is therefore very difficult to have large groups of non-contributors to evolve into cooperative behaviour, even if punishment is allowed. As illustrated in Figures 3 and 4, those groups are unlikely to develop sufficient pro-social agents to establish a stable population of co-operators who are willing to enforce cooperation.

Figure 6 shows the evolved values of α and β. We observed substantial differences between simulations with and without punishment. The evolved average behavioral type varied with increasing group sizes. For instance, in a group size of 2, agents who valued the earnings of others at a lower level than their own emerged in simulations with (and without) punishment. With a group size of 2 agents cooperate due to direct reciprocity forces. For group sizes 4 and 8 the agents who experience an evolution without punishment evolve norms with greater weight on the earnings of others. Groups who have more individuals with higher values of α and β are more likely to win intergroup conflicts. In Figure 7 we see the distributions of α and β at the end of the simulations for different group sizes. With group sizes 4 and 8 the evolved agents are mainly in the corner of high α's and β's. For group sizes of 16 and higher the evolved agents are highly concentrated in the corner of competitive agents. If cultural group selection does not lead to cooperative outcomes, competition within the groups affects the survival of pro-social agents.

When punishment is allowed, there is a higher level of cooperation, but this is not due to a higher level of pro-social preferences of the agents. As Figure 4 illustrates, pro-social behaviour initially evolve, but once punishment and cooperative behaviour is established, competitive agents evolve. Competitive agents gain utility from earning more than others, which can be derived by punishing defectors. The main concentration of agent types is in the corner of α=-1 and β= -1, but there is a large diversity of other types co-existing (Figure 8). The threat of punishment and the expectation of cooperation leads to a stable cooperative outcome. For larger groups, these expectations do not evolve.

We will now report on some sensitivity analysis. Boyd et al. [7] showed the sensitivity of the probability of inter group conflict, and the probability of imitating an agent from another group. We hypothesized that higher levels of group competition (a higher value of conflict with other groups) lead to more benefits for cooperative groups and a higher level of cooperative agents. Figures 9 and 10 show that this indeed the case. A higher level of conflict rates leads to cooperative behavior for groups of 16 agents versus 8 agents (no punishment), but the effect for simulations with punishment is modest.

Each generation an agent considers imitating the behavior of one other agent. With a probability (1-*m*) an agent from her group is drawn and with a probability *m* an agent is drawn from an outside group. Generally speaking, increasing group mixing rates decreased cooperation levels since high-payoff defection strategies spread more quickly. Figures 11 and 12 show that this is indeed the case in our model. When agents cannot punish, cooperation is abysmally small even with large mixing rates at a group size of 8. When punishment is allowed, there is some modest level of cooperation for group size 128 even when the mixing rate is very low.

Discussion

Experiments in behavioral economics show that humans make decisions in social dilemmas based on expectations of what others will do. Those experiments also show that humans generally take into account the earnings of others to evaluate the utility of material payoffs [15]. The inclusion of insights from behavioral economics into a model of cultural group selection of cooperation leads to different levels of cooperation. Cooperation levels are higher in small groups (i.e., 8 or fewer agents). The primary driver at these group sizes is a combination of direct reciprocity and costly punishment. Furthermore, we find a comparatively higher level of cooperation even without punishment compared to the results of Boyd et al. [7]. This contrast with the Boyd et al. [7] model is likely due to the influence of social preferences in our agents that promote equitable payoffs within groups.

Cooperation levels drop off quite rapidly in larger groups. This sharp decline of cooperation is caused by the challenge of groups without co-operators to establish sufficient trust to derive cooperation. With low levels of trust only pro-social agents cooperate. The larger the group size, the more pro-social agents need to be to derive cooperation. Larger groups are consequently also more sensitive to mistakes and invasions of defective behavior. Since decision making is not only based on behavioral types, but also on the expectation of the behavior by others, we observe a qualitatively different relationship between cooperation levels and group size compared to Boyd et al. [7].

We also observe that higher levels of cooperation in simulations with the option of punishment are not caused by higher levels of pro-social preferences among the agents. As speculated by Jensen [21] peer-punishment can lead to the evolution of highly competitive agents. Initially the population establishes a high level of cooperation and a credible threat (high T_P), but subsequently agents who prey on mistakes and defection will derive an evolutionary advantage. Hence, introducing the option of punishment allows different types of agent behavior to evolve.

Our analysis shows some of the consequences of including more behavioral complexity in agent-based models of social dilemmas. However, how do our results relate with the high levels of cooperation we observe in large human societies? Conditional cooperation is sensitive to the erosion of trust in groups, which is difficult to regain once lost. This may explain why complex human societies are often characterized by different kinds of rituals, symbols, and social norms that may be especially useful for developing trust relationships. As the late Elinor Ostrom emphasized in her work, trust is the five letter word that is key to keep complex societies functioning [27]. Others have recognized that human societies might be best represented as individual ecologies of games [6]. Experimental research has shown that the alteration of public good games with a reciprocity game leads to a higher level of cooperation compared with playing only public good games [23]. Field work on the study of governance of the commons has shown that effective communities use conflict resolution mechanisms as a way to restore trust relationships [26].

Peer punishment leads to spite and highly competitive agents [21]. To derive effective punishment in large group without this dark side of cooperation, institutionalized third party punishment is an option [8]. In such an arrangement, agents contribute to a public good which execute monitoring and enforcement.

By formulating agents as conditional strategists in a cultural group selection model, we find that cooperation is sensitive to the expectations of the behavior of others. In real societies, repeated activities are instituted to maintain and develop trust relationships. Future work may focus on the trade-offs in the costs of maintenance of social capital, third party punishment arrangements and the evolution of cooperative behavior.

Acknowledgement: We appreciate the support of the National Science Foundation for the grant BCS 0748632 and the comments of two anonymous reviewers.

References

[1] Alger, I., and Weibull, J.W., Homo Moralis – Preference evolution under incomplete information and assortative matching, *Econometrica* **81**(6) (2013): 2269–2302.

[2] Andreoni, J., and Miller, J.H., Giving according to GARP: An experimental study of rationality and altruism, *Econometrica* **70** (2002) 737-753.

[3] Axelrod, R., *The Evolution of Cooperation* (Basic Books, New York, 1984)

[4] Axelrod, R., *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton University Press, NJ., 1997)

[5] Alexander, R. D., *The Biology of Moral Systems* (Aldine de Gruyter, New York, 1987)

[6] Bednar, J., and Page, S., Can Game(s) Theory Explain Culture? The Emergence of Cultural Behavior Within Multiple Games. *Rationality and Society* **19**(1) (2007) 65–97.

[7] Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J., The evolution of altruistic punishment, *Proceedings of the National Academy of Science of the United States of America* **100**(6) (2003) 3531-3535.

[8] Boyd, R., Gintis, H., and Bowles, S., Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rate, *Science* **328** (5978) (2010) 617-620.

[9] Camerer, C., *Behavioral game theory: experiments in strategic interaction* (Princeton University Press, NJ, 2003).

[10] Charness, G., and Rabin, M., Understanding social preferences with simple tests. *Quarterly Journal of Economics* **117** (2002) 817-869.

[11] Chaudhuri, A., Sustaining Cooperation in Laboratory Public Goods Experiments: a Selective Survey of the Literature. *Experimental Economics*, 14(1) (2011): 47 – 83.

[12] Denant-Boemont L, Masclet D, and Noussair C., Punishment, counterpunishment and sanction enforcement in a social dilemma experiment, *Economic Theory* **33** (2007) 145–167.

[13] Fehr, E., and Schmidt, K., A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics* **114** (1999) 817-868.

[14] Fehr, E., and Gächter, S., Altruistic punishment in humans, *Nature* **415** (2002) 137-140.

[15] Fischbacher, U., Gächter, S., and Fehr, E., Are people conditionally cooperative? Evidence from a public goods experiment, *Economics Letters* **71**(3) (2001): 397-404

[16] Gintis, H., *Game theory evolving; a problem-centered introduction to modelling strategic interaction* (Princeton University Press, NJ, 2000).

[17] Güth, W., and Yaari, M., An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game, *In Explaining Process and Change—Approaches to Evolutionary Economics*, ed. by U. Witt. (University of Michigan Press, Ann Arbor 1992), pp. 23-34.

[18] Hamilton, W. D., Genetical evolution of social behavior I and II. *Journal of Theoretical Biology* **7** (1964) 1-52.

[19] Janssen, M.A. and Bushman, C., Evolution of cooperation and altruistic punishment when retaliation is possible, *Journal of Theoretical Biology* **254**: (2008) 541-545.

[20] Janssen, M.A., Holahan, R., Lee, A., and Ostrom, E., Lab Experiments for the Study of Social-Ecological Systems, *Science* **328** (2010) 613-617.

[21] Jensen, K., Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B* **365** (2010) 2635–2650.

[22] Milinski, M., Semmann, D., and Krambeck, H.J., Reputation helps solve the 'tragedy of the commons', *Nature* **415** (2002) 424-426.

[23] Nikiforakis, N., Feedback, punishment and cooperation in public good experiments. Games and Economic Behavior 68(2) (2010): 689-702

[24] Nowak, M.A. Five Rules for the evolution of cooperation, *Science* **314** (2006) 1560-1563.

[25] Nowak, M. A., and Sigmund, K., Evolution of indirect reciprocity by image scoring. *Nature* **393** (1998) 573-577.

[26] Ostrom, E., *Governing the commons* (Cambridge University Press, MA, 1990).

[27] Ostrom, E., Beyond Markets and States: Polycentric Governance of Complex Economic Systems, *American Economic Review* **100** (2010) 641-672.

[28] Ostrom, E., Walker, J., and Gardner, R., Covenants with and without a sword: self-governance is possible. *American Political Science Review* **86**(2) (1992) 404-417.

[29] Rabin, M., Incorporating fairness into game theory and economics, *American Economic Review* **83** (1993) 1281-1302.

[30] Rand D., and Nowak M., The evolution of antisocial punishment in optional public goods games, *Nature Communications* **2** (2011) 434.

[31] Trivers, R., The evolution of reciprocal altruism. *Quarterly Review of Biology* **46** (1971) 35-57.

[32] Yamagishi, T., The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* **51**(1) (1986) 110-116.

[33] Zahavi, A., The cost of honesty. *Journal of Theoretical Biology* **67** (1977) 603-605.

Figure captions

Figure 1: Conditions for agents to decide to contribute or not to the public good when punishment was not possible. The four figures are defined for different levels of trust (T_C). The shaded area represents the combinations of α and β leading to contributions. Different shadings refer to different group sizes (2, 4, 8, 16, 32).

Figure 2: Conditions for agents to decide to contribute or not to the public good, and when to use punishment. The four figures are defined for different levels of trust (T_C) and expected level co-operators punish (T_P). The light shaded area represents the combinations of α and β leading to contributions, and the dark shared area represents agent types who also use punishment. The lines indicate different group sizes (2, 4, 8, 16, 32, 64) above which agents cooperate.

Figure 3. The average level of cooperation for a typical run of group size 8 with and without punishment for a duration of 5000 generations.

Figure 4. The average levels of α and β of the agents for a typical run of group size 8 with and without punishment for a duration of 5000 generations.

Figure 5: The average level of cooperation for 50 runs for different group sizes *n*. Results with a replication of the Boyd et al. [7] model are compared with the model presented in this paper. For both models, two treatments are distinguished: no punishment and punishment.

Figure 6: Average level of α and β for two conditions (with and without punishment), for different group sizes.

Figure 7: Distribution of evolved agents for different levels of α and β when punishment was not possible. The darker the color the higher the frequency of observed agents at the end of the simulations. The figures are derived for group sizes 4, 8, 16 and 32.

Figure 8: Distribution of evolved agents for different levels of α and β when punishment was possible. The darker the color the higher the frequency of observed agents at the end of the simulations. The figures are derived for group sizes 4, 8, 16 and 32.

Figure 9. The evolved levels of cooperation with different levels of conflict rates (ex), and when there is no punishment.

Figure 10. The evolved levels of cooperation with different levels of conflict rates (ex), and when there is punishment.

Figure 11. The evolved levels of cooperation with different levels of mixing rates (m), and when there is no punishment.

Figure 12. The evolved levels of cooperation with different levels of mixing rates (m), and when there is punishment.